Matricial Norms and the Zeros of Polynomials

EMERIC DEUTSCH
Polytechnic Institute of Brooklyn
Brooklyn, New York

Communicated by Olga Taussky Todd

A matricial norm [1] is a mapping μ from the algebra M_n of complex $n \times n$ matrices into the set M_k^+ of nonnegative $k \times k$ matrices and which satisfies the following axioms:

(i) $\mu(\alpha A) = |\alpha|\mu(A) \qquad \forall \alpha \in C, \qquad \forall A \in M_n;$
(ii) $\mu(A + B) \leqslant \mu(A) + \mu(B) \qquad \forall A, B \in M_n;$
(iii) $\mu(AB) \leqslant \mu(A)\mu(B) \qquad \forall A, B \in M_n;$
(iv) $\mu(A) \neq 0$ if $A \neq 0.$

Here C denotes the complex field. The set M_k^+ is partially ordered componentwise, i.e., $(a_{ij}) \leqslant (b_{ij})$ if and only if $a_{ij} \leqslant b_{ij}$ for all $i, j = 1, \ldots, n$. If $k = 1$, then μ is a matrix norm [2]. Denoting by $r(A)$ the spectral radius of an $n \times n$ matrix A, it has been proved [1, 5] that

$$r(A) \leqslant r(\mu(A)), \tag{1}$$

which generalizes a well-known property of matrix norms [2].

A particular class of matricial norms can be obtained in the following manner. For an arbitrary complex $p \times q$ matrix $B = (b_{ij})$ denote

$$\psi(B) = \max_{i=1,\ldots,p} (|b_{i1}| + |b_{i2}| + \cdots + |b_{iq}|);$$

that is, $\psi(B)$ is the row norm of B. If

$$A = \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1k} \\ A_{21} & A_{22} & \ldots & A_{2k} \\ & & \vdots & \\ A_{k1} & A_{k2} & \ldots & A_{kk} \end{bmatrix}$$

is an arbitrary but fixed partitioning of the $n \times n$ matrix A, where $A_{11}, A_{22}, \ldots, A_{kk}$ are square matrices, then the mapping

$$\phi: M_n \to M_k^+,$$

$$\phi(A) = (\psi(A_{ij}))_{i,j=1,\ldots,k} \qquad (A \in M_n) \tag{2}$$

is a matricial norm on M_n [1].

In this paper we will apply matricial norms to obtain upper bounds for the zeros of the polynomial

$$f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0,$$

where $a_j \in C$ $(j = 0, 1, \ldots, n-1)$. We will denote by $\rho(f)$ the largest of the absolute values of the zeros of $f(z)$.

It is known that the zeros of $f(z)$ are the eigenvalues of its companion matrix:

$$F = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & 0 & -a_2 \\ & & & \vdots & & & \\ 0 & 0 & 0 & \ldots & 1 & 0 & -a_{n-2} \\ 0 & 0 & 0 & \ldots & 0 & 1 & -a_{n-1} \end{bmatrix}.$$

Let $k_0, k_1, \ldots, k_{n-2}$ be arbitrary positive numbers and denote $D = \operatorname{diag}(k_0, k_1, \ldots, k_{n-2}, 1) \in M_n$. The matrix

$$D^{-1}FD = \left[\begin{array}{cccccc|c} 0 & 0 & 0 & \ldots & 0 & 0 & -\dfrac{a_0}{k_0} \\ \dfrac{k_0}{k_1} & 0 & 0 & \ldots & 0 & 0 & -\dfrac{a_1}{k_1} \\ 0 & \dfrac{k_1}{k_2} & 0 & \ldots & 0 & 0 & -\dfrac{a_2}{k_2} \\ & & & \vdots & & & \\ 0 & 0 & 0 & \ldots & \dfrac{k_{n-3}}{k_{n-2}} & 0 & -\dfrac{a_{n-2}}{k_{n-2}} \\ \hline 0 & 0 & 0 & \ldots & 0 & k_{n-2} & -a_{n-1} \end{array}\right] \tag{3}$$

has the same eigenvalues as F, and so

$$\rho(f) = r(F) = r(D^{-1}FD). \tag{4}$$

Let $\phi\colon M_n \to M_2{}^+$ denote the matricial norm given by (2) corresponding to the partitioning shown in (3). Then

$$\phi(D^{-1}FD) = \begin{bmatrix} \beta & \gamma \\ k_{n-2} & |a_{n-1}| \end{bmatrix}, \tag{5}$$

where

$$\beta = \max\left\{\frac{k_0}{k_1}, \frac{k_1}{k_2}, \ldots, \frac{k_{n-3}}{k_{n-2}}\right\}, \tag{6}$$

$$\gamma = \max\left\{\frac{|a_0|}{k_0}, \frac{|a_1|}{k_1}, \ldots, \frac{|a_{n-2}|}{k_{n-2}}\right\}. \tag{7}$$

Applying inequality (1) to $D^{-1}FD$, we obtain from (4)

$$\rho(f) \leqslant r(\phi(D^{-1}FD)). \tag{8}$$

If σ is an arbitrary matrix norm on M_2, we obtain from (8)

$$\rho(f) \leqslant \sigma(\phi(D^{-1}FD)). \tag{9}$$

Obviously, the upper bound given by (9) cannot be better than the one given by (8). We will take for σ either the column norm σ_1 or the Euclidean norm σ_2 defined, respectively, by

$$\sigma_1(A) = \max\{|a_{11}| + |a_{21}|, |a_{12}| + |a_{22}|\},$$

$$\sigma_2(A) = (|a_{11}|^2 + |a_{12}|^2 + |a_{21}|^2 + |a_{22}|^2)^{1/2},$$

where $A = (a_{ij})_{i,j=1,2}$. Evaluating the right-hand sides of (8) and (9), we obtain

PROPOSITION 1. If $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, and k_0, $k_1, \ldots, k_{n-2}$ are arbitrary positive numbers, then

$$\rho(f) \leqslant \tfrac{1}{2}[\beta + |a_{n-1}| + \sqrt{(|a_{n-1}| - \beta)^2 + 4\gamma k_{n-2}}], \tag{10}$$

$$\rho(f) \leqslant \max\{\beta + k_{n-2}, \gamma + |a_{n-1}|\}, \tag{11}$$

$$\rho(f) \leqslant \sqrt{\beta^2 + \gamma^2 + k_{n-2}^2 + |a_{n-1}|^2}, \tag{12}$$

where β and γ are given by (6) and (7), respectively.

Remark 1. Since inequalities (11) and (12) have been obtained by applying a matrix norm to $\phi(D^{-1}FD)$, they cannot give better bounds than (10), of which the right-hand side is the spectral radius of $\phi(D^{-1}FD)$. We will see in Example 1 that even inequality (11) can yield a better upper bound for $\rho(f)$ than the following known inequality [6]:

$$\rho(f) \leqslant \max\left\{\frac{|a_0|}{k_0}, \frac{k_0 + |a_1|}{k_1}, \frac{k_1 + |a_2|}{k_2}, \ldots, \frac{k_{n-3} + |a_{n-2}|}{k_{n-2}}, k_{n-2} + |a_{n-1}|\right\}. \tag{13}$$

Incidentally, inequality (13) can be obtained by taking the row norm of the matrix $D^{-1}FD$.

Taking $k_0 = k_1 = \cdots = k_{n-2} = 1$, we obtain from Proposition 1

COROLLARY 1. *If* $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, *then*

$$\rho(f) \leqslant \tfrac{1}{2}[1 + |a_{n-1}| + \sqrt{(a_{n-1} - 1)^2 + 4M}], \tag{14}$$

$$\rho(f) \leqslant \max\{2, |a_0| + |a_{n-1}|, |a_1| + |a_{n-1}|, \ldots, |a_{n-2}| + |a_{n-1}|\}, \tag{15}$$

$$\rho(f) \leqslant \sqrt{2 + M^2 + |a_{n-1}|^2}, \tag{16}$$

where $M = \max\{|a_0|, |a_1|, \ldots, |a_{n-2}|\}$.

Remark 2. Inequality (14) has been proved in a different way in [3].

Remark 3. Even the upper bound given by (15) can be better than Cauchy's [4]:

$$\rho(f) \leqslant \max\{|a_0|, 1 + |a_1|, 1 + |a_2|, \ldots, 1 + |a_{n-1}|\}. \tag{17}$$

The latter is obtained from (13) by taking $k_j = 1$ ($j = 0, 1, \ldots, n-2$).

Example 1. Consider $f(z) = z^3 - 0.5z^2 - 2z - 2$. Inequality (17) gives $\rho(f) \leqslant 3$, while (14), (15), and (16) give the upper bounds 2.19, 2.5, and 2.5, respectively.

Assuming that $a_j \neq 0$ for all $j = 0, 1, \ldots, n-1$, and taking $k_j = a_{j+1}$ ($j = 0, 1, \ldots, n-2$) in Proposition 1, we obtain

COROLLARY 2. *If* $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ *with* $a_j \neq 0$ $(j = 0, 1, \ldots, n-1)$, *then*

$$\rho(f) \leqslant \tfrac{1}{2}[\beta' + |a_{n-1}| + \sqrt{(|a_{n-1}| - \beta')^2 + 4\gamma'|a_{n-1}|}], \tag{18}$$

$$\rho(f) \leqslant |a_{n-1}| + \max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_1}{a_2}\right|, \ldots, \left|\frac{a_{n-2}}{a_{n-1}}\right|\right\}, \tag{19}$$

$$\rho(f) \leqslant \sqrt{2|a_{n-1}|^2 + \beta'^2 + \gamma'^2}, \tag{20}$$

where

$$\beta' = \max\left\{\left|\frac{a_1}{a_2}\right|, \left|\frac{a_2}{a_3}\right|, \ldots, \left|\frac{a_{n-2}}{a_{n-1}}\right|\right\},$$

$$\gamma' = \max\left\{\left|\frac{a_0}{a_1}\right|, \left|\frac{a_1}{a_2}\right|, \ldots, \left|\frac{a_{n-2}}{a_{n-1}}\right|\right\}.$$

Remark 4. Corollary 2 can be modified for the case when some of the a_j's are equal to zero. Then, in Proposition 1, one should take $k_j = a_{j+1}$ if $a_{j+1} \neq 0$, and $k_j = 1$ if $a_{j+1} = 0$.

Remark 5. If $\beta' = \gamma'$, then the right-hand sides of (18) and (19) coincide.

Remark 6. Inequality (19) can yield a better upper bound for $\rho(f)$ than Kojima's [4]:

$$\rho(f) \leqslant \max\left\{\left|\frac{a_0}{a_1}\right|, 2\left|\frac{a_1}{a_2}\right|, 2\left|\frac{a_2}{a_3}\right|, \ldots, 2\left|\frac{a_{n-2}}{a_{n-1}}\right|, 2|a_{n-1}|\right\}. \tag{21}$$

The latter is obtained from (13) by taking $k_j = |a_{j+1}|$ $(j = 0, 1, \ldots, n-2)$.

Example 2. Consider $f(z) = z^4 - 4z^3 + 3z^2 + 2z - 1$. Inequality (21) gives $\rho(f) \leqslant 8$, while both (18) and (19) give $\rho(f) \leqslant 4.75$.

Let t be an arbitrary positive number and in Proposition 1 take $k_j = t^{n-j-1}$ $(j = 0, 1, \ldots, n-2)$. We obtain

PROPOSITION 2. If $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, and t is an arbitrary positive number, then

$$\rho(f) \leqslant \tfrac{1}{2}[t + |a_{n-1}| + \sqrt{(|a_{n-1}| - t)^2 + 4\delta t}], \tag{22}$$

$$\rho(f) \leqslant \max\left\{2t, \frac{|a_0|}{t^{n-1}} + a_{n-1}, \frac{|a_1|}{t^{n-2}} + a_{n-1}, \ldots, \frac{|a_{n-2}|}{t} + |a_{n-1}|\right\}, \tag{23}$$

$$\rho(f) \leqslant \sqrt{2t^2 + \delta^2 + |a_{n-1}|^2}, \tag{24}$$

where

$$\delta = \max\left\{\frac{|a_0|}{t^{n-1}}, \frac{|a_1|}{t^{n-2}}, \ldots, \frac{|a_{n-2}|}{t}\right\}.$$

Remark 7. Inequality (23) can yield a better upper bound for $\rho(f)$ than Wilf's [6]:

$$\rho(f) \leqslant \max\left\{\frac{|a_0|}{t^{n-1}}, t + \frac{|a_1|}{t^{n-2}}, t + \frac{|a_2|}{t^{n-3}}, \ldots, t + \frac{|a_{n-2}|}{t}, t + |a_{n-1}|\right\}. \tag{25}$$

The latter is obtained from (13) by taking $k_j = t^{n-j-1}$ ($j = 0, 1, \ldots, n-2$).

Example 3. Consider $f(z) = z^4 + 2z^3 - 13z^2 - 38z - 24$. Inequality (25) gives $\rho(f) \leqslant 7.34$, while (22) and (23) give $\rho(f) \leqslant 6.14$ and $\rho(f) \leqslant 6.34$, respectively.

Assuming that $a_{n-1} \neq 0$ and, in Proposition 2, taking $t = |a_{n-1}|$, we obtain

COROLLARY 3. *If* $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$ *with* $a_{n-1} \neq 0$, *then*

$$\rho(f) \leqslant |a_{n-1}| + (\delta')^{1/2}, \tag{26}$$

$$\rho(f) \leqslant |a_{n-1}| + \max\left\{|a_{n-1}|, \frac{\delta'}{|a_{n-1}|}\right\}, \tag{27}$$

$$\rho(f) \leqslant \sqrt{3|a_{n-1}|^2 + \delta'^2/|a_{n-1}|^2}, \tag{28}$$

where

$$\delta' = \max\left\{\frac{|a_0|}{|a_{n-1}|^{n-2}}, \frac{|a_1|}{|a_{n-1}|^{n-3}}, \ldots, \frac{|a_{n-3}|}{|a_{n-1}|}, |a_{n-2}|\right\}.$$

Denote

$$N = \max\{|a_0|^{1/n}, |a_1|^{1/(n-1)}, \ldots, |a_{n-3}|^{1/3}, |a_{n-2}|^{1/2}\}. \tag{29}$$

Then $|a_j| \leqslant N^{n-j}$ ($j = 0, 1, \ldots, n-2$) with equality for at least one j. Taking, in Proposition 2, $t = N$, we obtain

COROLLARY 4. *If* $f(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, *then*

$$\rho(f) \leqslant \tfrac{1}{2}[N + |a_{n-1}| + \sqrt{(|a_{n-1}| - N)^2 + 4N^2}], \tag{30}$$

$$\rho(f) \leqslant N + \max\{N, |a_{n-1}|\}, \tag{31}$$

$$\rho(f) \leqslant \sqrt{3N^2 + |a_{n-1}|^2}, \tag{32}$$

where N *is given by* (29).

Remark 8. From inequality (31) we obtain Fujiwara's upper bound [4]:

$$\rho(f) \leqslant 2 \max_{j=1,\ldots,n} |a_{n-j}|^{1/j},$$

which is weaker than both (30) and (31) (see Remark 1).

REFERENCES

1 E. Deutsch, Matricial norms, *Numer. Math.* **16**(1970) 73–84.

2 A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York-Toronto-London, 1964.

3 A. Joyal, G. Labelle, and Q. I. Rahman, On the location of zeros of polynomials, *Canad. Math. Bull.* **10**(1967), 53–63.

4 M. Marden, The geometry of the zeros of a polynomial in a complex variable, *Mathematical Surveys*, Amer. Math. Soc., 1949.

5 F. Robert, Etude et utilisation de normes vectorielles en analyse numérique linéaire, Thèsis, Université de Grenoble, 1968.

6 H. S. Wilf, Perron-Frobenius theory and the zeros of polynomials, *Proc. Amer. Math. Soc.* **12**(1961), 247–250.

Received March, 1970